

# Small-Area Spatiotemporal Analysis of Pedestrian and Bicyclist Injuries in New York City 2001-2010

August 27, 2014

## Abstract

**Objective:** This study quantifies the spatiotemporal risk of pedestrian and bicyclist injury in New York City at the census tract level over a recent 10-year period, identifies areas of increased risk, and evaluates the role of socioeconomic and traffic-related variables in injury risk.

**Methods:** Crash data on 140,835 pedestrian and bicyclist injuries in 1,908 census tracts from 2001 to 2010 were obtained from the New York City Department of Transportation. Injury counts within census tracts were analyzed with Bayesian hierarchical spatial models using Integrated Nested Laplace Approximations (INLA). The model included variables for social fragmentation, median household income, average vehicle speed and traffic density, a spatially unstructured random effect term, a spatially structured conditional autoregression CAR term, a first-order random walk correlated time variable, and an interaction term for time and place. Incidence Density Ratios (IDR), credible intervals (CrI) and probability exceedances were calculated and mapped.

**Results:** The yearly rate of pedestrian crashes decreased 16.2% over the study period, from 23.7 per 10,000 population to 16.2 per 10,000. The temporal term in the spatiotemporal model indicated that much of the decrease over the study period occurred during the first four years of the study period. Despite an overall decrease, the model identified census tracts that were at persistently increased risk of pedestrian injury throughout the study period, as well as areas that experienced sporadic annual increases in risk. Aggregate social, economic and traffic-related measures were associated with pedestrian injury risk at the ecologic level. Every one unit increase in a standardized social fragmentation index was associated with a 19% increase in pedestrian injury risk (IDR = 1.19, 95% CrI 1.16, 1.23), and every one standardized unit increase in traffic density was associated with a 20% increase in pedestrian injury risk (IDR=1.20, 95% CrI 1.15, 1.26). Each 10 mile per hour increase in average traffic speed in a census tract was associated with a 24% decrease in pedestrian injury risk (IDR=1.24, 95% CrI 0.69, 0.83).

**Conclusions:** The risk of a pedestrian or bicyclist being struck by a motor vehicle in New York

City decreased from 2000 to 2004. Some census tracts in the city did not benefit from over-all reductions, or experienced sporadic years of increased risk compared to the city as a whole. Injury risk at the census-tract level was associated with social, economic and traffic-related factors.

## Introduction

After many years of notable declines, the number of pedestrian and bicyclist fatalities across the United States (US) has increased nearly 5% between 2009 and 2011.<sup>1,2</sup> This is in contrast to continued declines in both rates and frequencies of motor vehicle occupant fatalities, and is reflected in the increasing proportion of traffic fatalities due to pedestrian injury. Walking accounts for 2.8% of all commutes and 8.6% of all trips in the US, but pedestrians represented 11.3% (4,645) of the 41,059 total US traffic fatalities in 2007.<sup>3,4</sup>

Pedestrian injury and fatality is particularly pernicious in dense urban environments. In 2007 73% of pedestrian crashes in the US occurred in urban areas.<sup>4</sup> In New York City, pedestrian deaths have outnumbered occupant deaths since 1910, and between 1997 and 2006, New York City accounted for the largest proportion of pedestrian fatalities in the United States.<sup>5</sup>

The City of New York has in recent years undertaken interventions to help insure the safety of all road users. Efforts include the installation of 1,500 pedestrian signals, re-engineering 60 miles of streets and 20 intersections, the evaluation 20 mile-per-hour pedestrian safety zones, and piloting a program to improve left-turn visibility in Manhattan. Total traffic fatalities in New York City decreased 35% between 2001 and 2009, making 2009 the safest year on record in terms of traffic crashes.<sup>6</sup> More recently, the office of the mayor has instituted a "Vision Zero" traffic plan to address the "epidemic of traffic fatalities and injuries" in New York City.<sup>7</sup> The initiative is based on pioneering efforts in Sweden to bring traffic fatalities down to zero,<sup>8</sup> and focuses in large part on pedestrians and bicyclists. The New York City plan involves educational, enforcement, and engineering interventions, as well as research, surveillance and data analysis "to help target traffic safety interventions" and evaluate the effectiveness of the program.<sup>7</sup> (p13)

This study quantifies the spatiotemporal risk of pedestrian injury in New York City at the census tract level over a recent 10-year period using a Bayesian hierarchical spatial regression model<sup>9</sup> with Integrated Nested Laplace approximations.<sup>10,11</sup> The study presents city-wide and local tempo-

ral trends, measures local risk, and examines the association of social, economic and traffic-related variables with that risk. The primary goals of the study are to describe the spatiotemporal epidemiology of pedestrian and bicyclist injuries, define risk at the census tract level that can be used to measure the effectiveness of interventions, identify areas that require investigation and additional intervention, and establish an approach to pedestrian injury surveillance at the local level over place and time. Secondary goals are to quantify and interpret the association of ecologic-level variables that may contribute to or mediate pedestrian and bicyclist injury risk at the community level.

# Methods

## Data

Crash data were obtained from the New York City Department of Transportation, and consisted of police investigations for all motor-vehicle crashes in New York City for the years 2001 to 2010 involving personal injury or property damage in excess of \$1,000. Among the available variables were date, age and gender of the injured individuals, and GPS-based latitude and longitude of the crash location. The crash data were in a relational set of database files linked by a unique crash ID number, which were translated into a single flat text file with one observation for each pedestrian crash.

Pedestrians crashes were identified using an indicator variable for pedestrians ("Body Type = 35") or bicyclists ("Body Type=36") and were internally validated by comparing the crash ID number against a second variable in a separate linked table that also identified pedestrians ("Vehicle Type=6"). The data were read into the R statistical analysis program,<sup>12</sup> cleaned and evaluated for outliers, inconsistent values and missing entries. Date and time variables were translated into POSIX objects to extract time variables. Crash latitude and longitude coordinates were assigned to census tracts using the R maptools package.<sup>13</sup>

Census-tract populations were based on United States 2000 and 2010 decennial census enumerations.<sup>14</sup> Census tracts were restricted to 1929 census tracts that were present in both 2000 and 2010, creating a closed cohort of census tracts. Population estimates were linearly interpolated over inter-census years. Census tract housing and economic data were based on 2010 US Census estimates.

A social fragmentation index based on the work of Peter Congdon<sup>15,16</sup> was created using 4 variables extracted from US census variables: the proportion of total housing units in a census tract that were not owner occupied, the proportion of vacant housing units, the proportion of individuals living alone and the proportion of housing units into which an occupant recently moved. A "recent" move

in the Census data was defined as anytime after the prior census enumeration. The four component social fragmentation index variables were standardized and summed with equal weights, resulting in a normally distributed index with mean zero and 95% quantiles between -2.5 and 2.2.

An economic measure was based on the census tract median household income for the past 12 months (in units of 10,000 2012 inflation-adjusted dollars). Census tract traffic-related variables were obtained from the New York City Department of Health and Mental Hygiene, and consisted of traffic density (vehicle kilometers traveled per day per square kilometer) in standardized units, and average speed in increments of 10 miles per hour.<sup>17</sup>

A census tract map file for New York City was obtained from the US Census and was matched to aggregated census tract counts of pedestrian and bicyclist injuries. Twenty-one census tracts with zero population representing large parks, beaches, cemeteries and train yards were excluded from analysis. Three census tracts representing highly-trafficked tourist areas with small underlying populations (Grand Central, Penn Station, Times Square) were adjusted based on hotel capacity, adding 16,000 transient residents to each area.<sup>18</sup> A contiguous neighbor adjacency matrix graph was created for use with the conditional autoregression term in the model using the `poly2nb` function in the `spdep` R package.<sup>19</sup>

## **Statistical Modeling**

Counts of pedestrian and bicyclist injuries in New York City census tracts were spatially modeled according to Besag-York-Mollie as described by Lawson<sup>9,20,21</sup>

$$y_i \sim \text{Pois}(\lambda_i = e_i \theta_i)$$

$$\log(\theta_i) = \beta x_i + v_i + \eta_i$$

$$v \sim \text{nl}(0, \tau_v)$$

$$\eta \sim \text{nl}(\bar{\eta}_\delta, \tau_\eta/n_\delta)$$

where,

(1) the  $y_i$  counts in area  $i$ , are independently identically Poisson distributed and have an expectation in area  $i$  of  $e_i$ , the expected count, times  $\theta_i$ , the risk for area  $i$ .

(2) a logarithmic transformation ( $\log(\lambda_i)$ ) allows a linear, additive model of regression terms ( $\beta x_i$ ), along with

(3) a spatially *unstructured* random effects component ( $v_i$ ) that is i.i.d normally distributed with mean zero ( $\sim \text{nl}(0, \tau_\eta)$ ), and

(4) a conditional autoregressive spatially structured component<sup>22</sup> ( $\eta \sim \text{nl}(\bar{\eta}_\delta, \tau_\eta/n_\delta)$ ) in which a "neighborhood" consisting of spatially adjacent shapes is characterized by the normally distributed mean of the spatially structured random effect terms for the spatial shapes that make up the neighborhood ( $\bar{\eta}_\delta$ ), and the standard deviation of that mean divided by the number of spatial shapes in the neighborhood ( $\tau_\eta/n_\delta$ ). This spatially structured conditional autoregression component is also sometimes described as a Gaussian process  $\lambda \sim \text{NI}(W, \tau_\lambda)$  where  $W$  represents the matrix of neighbors that defines the neighborhood structure, and the conditional distribution of each  $\lambda_i$ , given all the other  $\lambda_i$  is normal with  $\mu$  = the average  $\lambda$  of it's neighbors and a precision ( $\tau_\lambda$ ).

The spatial model is extended to include temporal characteristics for a space-time model of outcome counts in small areas in fixed time periods. This approach extends the Besag-York-Mollie model to include a linear term for space-time interaction and a non-parametric spatiotemporal time



trend.<sup>23,24,25</sup> Space-time data can be thought of as multivariate or correlated observations of counts within fixed spatial and temporal periods, where, for  $i = 1 \dots M$  small areas and  $j = 1 \dots J$  time periods, an outcome,  $y_{ij}$ , is characterized by the risk  $\theta_{ij}$ .

A baseline convolution model that consisted solely of an intercept term with unstructured and spatially structured random effect terms was extended to include covariates for social fragmentation, median household income, traffic density and average speed. The convolution model with covariates was extended to a space-time interaction model. Candidate variables for space-time included uncorrelated normally distributed and correlated random walk terms for time in increments of year both alone and in combination with an interaction term for census tract. Model selection was based on deviance information criteria and number of effective parameters. Poisson, zero-inflated Poisson and negative binomial models were evaluated.

The final linear model consisted of an intercept ( $\beta_0$ ); a vector of census-level explanatory variables ( $\beta x_i^T$ ) for social fragmentation (Congdon index), economics (in \$10,000 increments of median household income), average speed (in 10-mile per hour increments), and traffic density (in standardized units); a spatially unstructured random effect term ( $v_i$ ); a spatially structured conditional autoregression term ( $\eta_i$ ); a first-order random walk correlated time variable ( $\gamma_j$ ), and an interaction term for time and place ( $\psi_{ij}$ ). The log of the yearly census-tract population was included as an offset variable.

$$\log(\theta_{ij}) = \beta_0 + \beta x_i^T + v_i + \eta_i + \gamma_j + \psi_{ij} + (offset)$$

The intercept is interpreted as the average city-wide risk on the log scale adjusted for the covariates, random effects and spatiotemporal terms. The exponentiated coefficients for the explanatory covariates are interpreted as incidence density ratios. The spatially unstructured random effect term captures normally-distributed or Gaussian random variation around the mean or intercept.

The spatially-structured conditional autoregression term accounts for local geographic influence. The space-time interaction is a random-effect term added to the linear model and is interpreted as the residual effect after the unstructured, spatially structured and time effects are modeled, and represents sporadic short-term outbreaks or clusters.

Spatial risk, controlling for or holding the covariates constant, was calculated as  $\zeta_i = v_i + \eta_i$ ,<sup>10</sup> and is interpreted as the residual spatial risk for each area (compared to the all of New York City) after social fragmentation, economics, average speed traffic density, and time trend are taken into account. The probability of spatial risk greater than 2 ( $\Pr[e^{\zeta_i} > 2]$ ) was calculated. As originally described by Clayton and Bernardinelli,<sup>26</sup> these exceedance probabilities are the posterior probabilities for an area's spatial risk estimate exceeding some pre-set value. This was extended by Richardson, et al<sup>27</sup> to decision rules "for classifying whether (an area) has an increased risk based on how much of the posterior distribution of the relative risk parameter ... exceeds a reference threshold".<sup>28</sup> They are calculated as the proportion of simulations for which the linear combination of effects ( $\zeta$ ) exceeds the target value. Lastly, the proportion of spatially explained variance was calculated as the proportion of total spatial heterogeneity accounted for by the spatially structured conditional autoregression variance.<sup>10</sup>

Spatial modeling was conducted using integrated nested Laplace approximations (INLA) with the R INLA package<sup>11</sup> using approaches described by Blangiardo, et al.<sup>10</sup> The study protocol was approved as exempt by the (*Deleted for Blind Peer Review*) institutional review board as IRB-AAAF3448.

## Results

There were 168,806 pedestrian and bicyclist injuries in New York City between 2001 to 2010, of which 1,153 (0.7%) were listed as "bicyclists" and 166,907 (99.3%) were listed as "pedestrians". For the remainder of this paper, pedestrian refers to both pedestrians and bicyclists. The mean age of an injured pedestrian in New York City between 2001 and 2010 was 34.6 years, with a median of 31 years. There were 76,737 females (49.3%) and 78,897 males (50.7%); 12,426 entries (7.4%) were missing data on gender. The population-based rate of pedestrian injuries decreased 16.2% over the study period, from 23.7 per 10,000 to 16.2 per 10,000. (Figure 1) Of the total 168,806 entries, 140,835 pedestrian crashes (83.4%) had sufficient information to be geocoded to one of the 1,929 census tracts within New York City that existed in both 2001 and 2010. Twenty one census tracts for which there were no population data were removed, and 3 highly trafficked but low population census tracts were adjusted. The final spatial analytic data set of 1,908 census tracts.

The results for the fixed effects of the covariates for the the space-time interaction model are presented in Table 1. Holding all other covariates to zero and adjusting for spatiotemporal variation, every one unit increase in the social fragmentation index was associated with a 19% increase in pedestrian injury risk (95% Credible Interval (CrI) 1.16, 1.23). Similarly, for every one standardized unit increase in traffic density, there was a 20% increase in pedestrian injury risk (95% CrI 1.15, 1.26). For every 10 mile per hour increase in average traffic speed in a census tract, there was a 24% decrease in pedestrian injury risk (95% CrI 0.69, 0.83). Each \$10,000 increment in median household income within a census tract was associated with a 3% decrease in census-tract level pedestrian injury risk (95% CrI 0.95, 0.99).

The spatially unstructured heterogeneity random effect term was normally distributed and spatially random. The spatially structured conditional autoregression term demonstrated areas of spatial patterning and similarity among census tracts. The spatial risk term at the census tract level identified areas at increased risk of pedestrian injury throughout the 10-year period. (Figure 2) Posterior probabilities for an area's spatial risk estimate exceeding 2 are presented in Figure 3. The proportion

of the total spatial heterogeneity explained by the spatially structured conditional autoregression term was 73.2%.

The temporal trend for the 10-year study period, holding the covariates and spatial risk constant, showed an initial steep decrease in the risk of pedestrian injury in New York City between 2001 and 2004 followed by a leveling off. (Figure 4) Probability exceedances for the yearly the space-time interactions are presented for 4 of the study period years, (Figure 5) and identify residual isolated areas of greater than twice the city-wide risk after the fixed covariate effects, unstructured, spatially structured and time random effects are held constant. The fewest such areas of increased risk are seen for 2004, when overall city-wide risk was at its lowest.

## Discussion

This work adds to the existing research in pedestrian injury epidemiology in three ways. The study presents pedestrian and bicyclist injury risk assessment at a finer level of geographic detail than has been generally been reported, which allows more opportunity for focused investigations and interventions. The methods in this study have previously not been used for pedestrian injury research and offer practical tools for spatial analysis in injury epidemiology that might not otherwise be readily apparent. And the study extends the well-recognized association of pedestrian injury with socioeconomic factors, to the construct of social fragmentation, which may offer additional insights into the complex interplay of physical and sociocultural environment in pedestrian injury risk.

Spatial clustering of pedestrian injuries in New York City is well recognized, with over 4 times as many fatal pedestrian injuries per mile of street occurring in the borough of Manhattan, compared to the other four boroughs,<sup>6</sup> and temporal trends of pedestrian injury in New York City have also been described.<sup>29</sup> This study seeks to add to these kinds of assessments by describing global and local patterns of pedestrian injury and quantifying the risk of injury at the census-tract level while controlling for socioeconomic, and vehicle-related characteristics like speed and traffic density. Combining time and place allowed the identification of areas for which there was evidence that the observed incidence density ratio of injury risk was greater than that for the city as a whole, given the underlying demographic and ecologic risks present in that place. These spatiotemporal analyses can be viewed as a kind of filter rolling across time and place. This kind of analysis can inform injury control initiatives like Vision Zero<sup>7</sup> by helping to identify areas that may require additional interventions and activities.

Holding temporal trends constant, place was an important determinant of pedestrian injury risk beyond other perhaps, more immediately apparent factors such as traffic density. There were areas in each borough of New York City, e.g. northern Manhattan, western Queens and central Brooklyn, where the risk of pedestrian injury was higher than the city as whole throughout the 10-year study

period. The use of exceedance probabilities refines this characterization of risk by placing explicit probabilities on the observed risk estimates to identify those areas for which the increased risk was highly unlikely to be due to chance. Alternatively, the temporal results of the study indicate that holding local risk constant, much of the decline in pedestrian injury risk in New York City during the study period occurred before 2004. The space-time interaction results identified areas of sporadic increased risk over the 10-year period.

Given the inherently ecological nature of this study, the parameter estimates cannot be interpreted causally. The results, do however, measure important social, cultural and contextual factors that may influence pedestrian injury risk, and perhaps lead to more proximal and modifiable factors that can be manipulated and changed to address that risk. It is increasingly apparent that the conditions in which people live, play and work have an influence on their health, including their risk of injury,<sup>30,31</sup> so the association of social fragmentation with pedestrian injury risk demonstrated in this study is not unexpected.

Despite its ecological nature, the results of studies of this kind may lead to fruitful interventions. Measures of social context similar to that used in this study have been linked to risk of suicide and self harm,<sup>15</sup> first admissions for psychosis<sup>32</sup> and myocardial infarction.<sup>33</sup> Social fragmentation has been associated with individual levels of walking and exercise that differ by gender,<sup>34</sup> and may directly influence urban pedestrian injury risk. Alternatively, such inherently ecological constructs as collective efficacy may point to similarly collective interventions. In a study of neighborhood cohesion and rates of violent crime,<sup>35</sup> the authors suggest that the effects are mediated through the "political economy". Perhaps political mobilization to effect change in the built environment leads to decreased pedestrian injury risk.

Pedestrian injury risk decreased with increasing average vehicle speed in a census tract. There are a number of possible explanations for this. We do not know the speed of the vehicles that struck pedestrians, so there is an element of ecologic bias at play. While the average speed was lower, the vehicles striking pedestrians may have been traveling at higher than average speeds.

In addition, slower speeds indicate more vehicles, which increases overall exposure to traffic. It may also be that because the average vehicle speed and traffic density variables in these analyses included highways, bridges and tunnels,<sup>17</sup> areas with higher average speeds would be less likely to be frequented by pedestrians, decreasing exposure. A sensitivity analysis using a version of the speed and density variables that did not include highways, bridges and tunnels, though, returned essentially the same results.

This study points out the importance of the intersection of time and location in epidemiologic analyses. Spatiotemporality may be the missing component in many epidemiologic inquiries. Epidemiologists routinely assess characteristics of persons (e.g., demographic characteristics and individual behavior), often consider geography (e.g. environmental exposures and infectious diseases), and increasingly consider time (follow-up periods, life-course analyses). Far less customarily though, do we consider how space and time combine to influence risk.<sup>36</sup> As epidemiologic inquiry proceeds to better locate individuals within their full place-lifetime continuum, and as we increasingly recognize the complex determination of disease,<sup>37,38,39</sup> understanding the intersecting role of space and time becomes central to the full understanding of certain health outcomes. Placing pedestrian injury into the context of time and place can be of immediate use to public health professionals and epidemiologists tasked with informing the efforts of traffic safety engineers, law enforcement officials, and policy makers implementing injury control and prevention programs.

Approaching public health data from a small-area spatial perspective poses challenges. The data are often aggregated to political or administrative groupings that may be too large and undifferentiated to capture risk appropriately.<sup>40,41</sup> Small-area analyses have developed into an effective and informative approach to address some of the challenges of spatial epidemiology.<sup>15</sup> Bayesian approaches are particularly useful when when data are indexed at fine geographical resolution.<sup>28</sup> Combined with the ubiquity of GIS and the increasing availability of spatiotemporal data, these methods can contribute to healthier safer environments.<sup>42</sup> In the field of injury epidemiology, Bayesian Markov Chain Monte Carlo (MCMC) methods have been used to model motor vehicle crash counts in a

way to elucidate the safety profile for different age groups.<sup>43</sup> Bayesian spatial hierarchical models are increasingly being applied to space-time crash data at the local level,<sup>44,45</sup> although traffic safety researchers report data and methodological constraints on the modeling and visualization of crash data at levels sufficient to discern important spatiotemporal patterns.<sup>44</sup>

Tools like INLA can address some practical constraints by providing less computationally expensive approaches to large, complex data sets at detailed geographic levels, using the kinds of random effect models often employed in spatial and spatiotemporal analyses. Unlike Markov Chain Monte Carlo methods that sample from the posterior distribution of parameters, INLA approximates the mode as the mean, and calculates the derivative in the area of the mode to approximate the variance. The closer the target posterior distribution is to Gaussian, the better the approximation.<sup>46</sup> While theoretically known for some time, Laplace approximations have only recently developed sufficiently to be considered accurate enough for practical application.<sup>10</sup> Other alternatives to MCMC are available,<sup>47</sup> but INLA is perhaps the most mature in terms of its default spatial analytic models and tools.<sup>10</sup>

The results for the social fragmentation index should be viewed cautiously. The index may not be as accurate in New York City as in other settings. New York City differs from other areas of the United States in ways that impact directly on the components that make up the index. New Yorkers are less likely to own their own homes (32% in New York vs. 65% national), the median value of a home is higher (\$500,000 vs \$180,000), and the population may be more mobile than other areas of the United States.<sup>14</sup> But, as noted previously, pedestrian injury is in many respects a uniquely urban phenomenon, and in that respect New York City resembles other US cities concerned enough about pedestrian injuries to contemplate Vision Zero initiatives. The proportion of home owners in San Francisco is about 36%, and in Boston it is about 34%.<sup>14</sup> Even in London, where the Congdon index was developed and first applied, home ownership is estimated at 50%.<sup>48</sup> And, in terms of reliance on walking and public transportation, New York City may resemble London more than most metropolitan areas in the United States. In fact, the first recorded motor vehicle crash fatalities



in history occurred in London and New York City in the 19th century, and were both pedestrians.<sup>49</sup>

The index and the traffic density variable are presented as standardized measures. Standardized measures in epidemiology have long had detractors<sup>50</sup> and defenders.<sup>51</sup> Importantly, such measures limit direct predictive interpretation and application by policy makers. In this study, the use of a standardized measure, while not inviting direct predictive interpretation, serves the purpose of directing attention to the complex interplay of physical and social environment and its possible association with pedestrian injury risk. Clinical epidemiology, and injury epidemiology in particular, is characterized by summary measures like the injury severity score that are useful because they are shown to correlate to morbidity and mortality, but do not inherently lend themselves to easy interpretation. Standardization also serves the second useful purpose of allowing comparisons, albeit imperfectly as note above, of constructs like social fragmentation across studies and populations. And, in this analysis, standardization allowed for better fit of the data to the model.

While the model represents an acceptable trade off between goodness of fit and complexity, efforts should include external validation of the results. This could be statistical, e.g. comparing predictions based on the model to actual data from years beyond 2010, or could involve statistical associations with external sources of covariate data. In one evaluation of fixed local covariates, the increased frequency of pedestrian crashes in low income, high minority areas in Chicago was shown to be associated with walkability and access to transit options.<sup>52</sup> Ideally, external validation would include the onsite evaluation of high-risk high-probability census tract areas. Site visits can yield deep and practical insights about local risks,<sup>53</sup> but are expensive, time consuming and require complex sampling strategies. In an intriguing alternative approach, virtual site visits could be conducted using tools like Google street view to collect and evaluate local geographic data with less expense and difficulty.<sup>54,55</sup>

Two of the statistical approaches used in this study come with caveats. First, exceedance probabilities, which have been proposed as a Bayesian approach to hotspot identification and are in relatively common use,<sup>28,56,57</sup> can be sensitive to model specifications.<sup>58</sup> And second, the propor-

tion of variance explained by the spatially structured conditional autoregression term is not strictly speaking a variance partition coefficient, because the structured and unstructured spatial terms may not be directly comparable. It is, though, an indication of the relative contribution of each of the spatial components.<sup>10</sup>

## **Conclusions**

The analyses in this study show that the risk of pedestrian injury in New York City decreased between 2000 to 2010, with much of the decline occurring during the first four years of the decade. Injury risk at the census tract level was associated with social, economic and traffic-related factors. Census tracts benefited unequally from the over-all reductions, with some areas having consistently elevated risks compared to the city-wide experience, and others experiencing sporadic years of increased risk. The kinds of small-area spatiotemporal Bayesian hierarchical modeling approaches used in this study are increasingly a practical option for epidemiologists interested in evaluating risk of disease outcomes in the context of place and time. Tools like INLA may allow the elucidation of the complex interplay of spacial and temporal risk at finer levels of aggregation.

## Tables and Figures

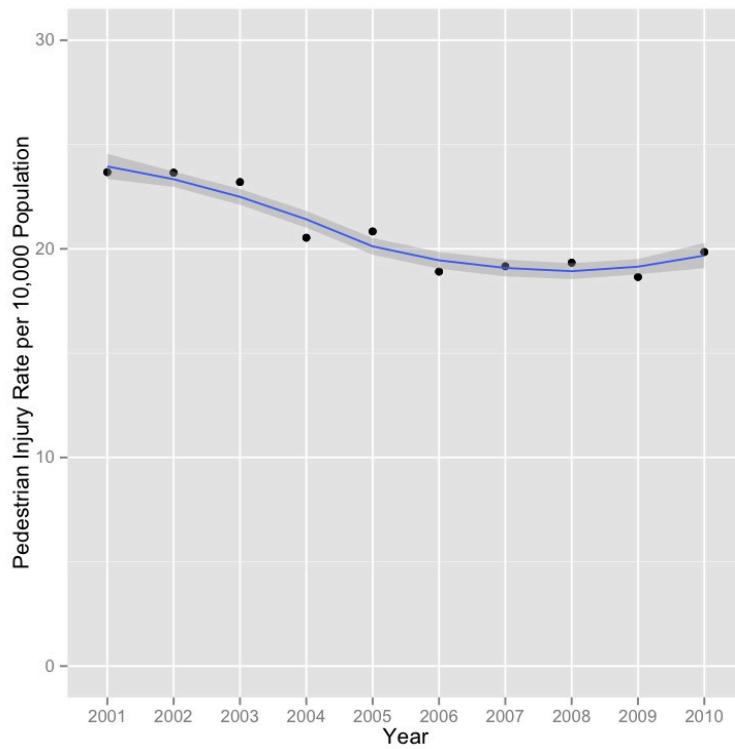


Figure 1: Rate of Pedestrian Injuries, All Ages, New York City, 2001 to 2010.

	Mean	0.025 Quantile	0.975 Quantile
Intercept	0.00	0.00	0.00
Social Fragmentation	1.19	1.16	1.23
Median Household Income	0.97	0.95	0.99
Traffic Density	1.20	1.15	1.26
Average Speed	0.76	0.69	0.83

Table 1: Results of Spatiotemporal Model of Pedestrian Injury Risk, New York City, 2001-2010

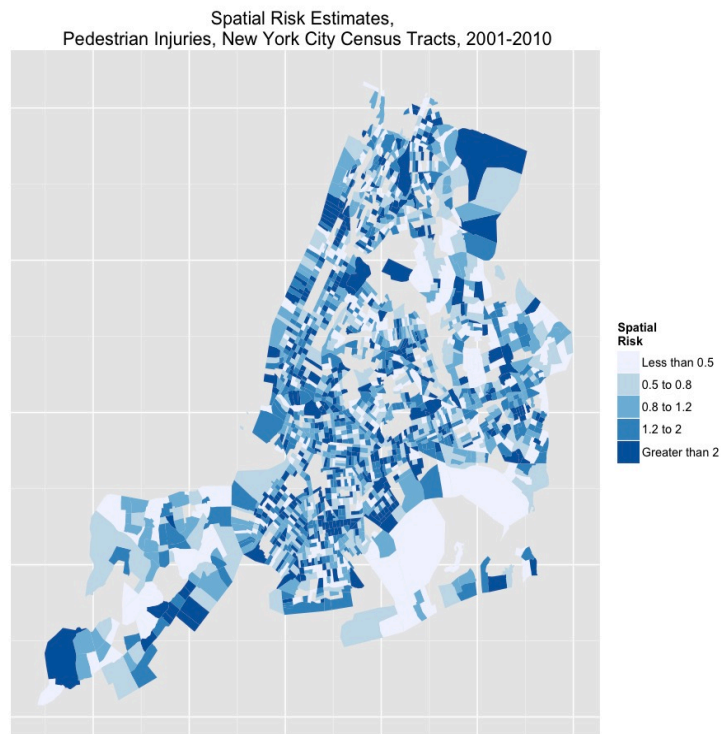


Figure 2: Incidence Density Ratios, Pedestrian Injury Risk, Spatiotemporal Model of New York City Census Tracts, 2001-2010

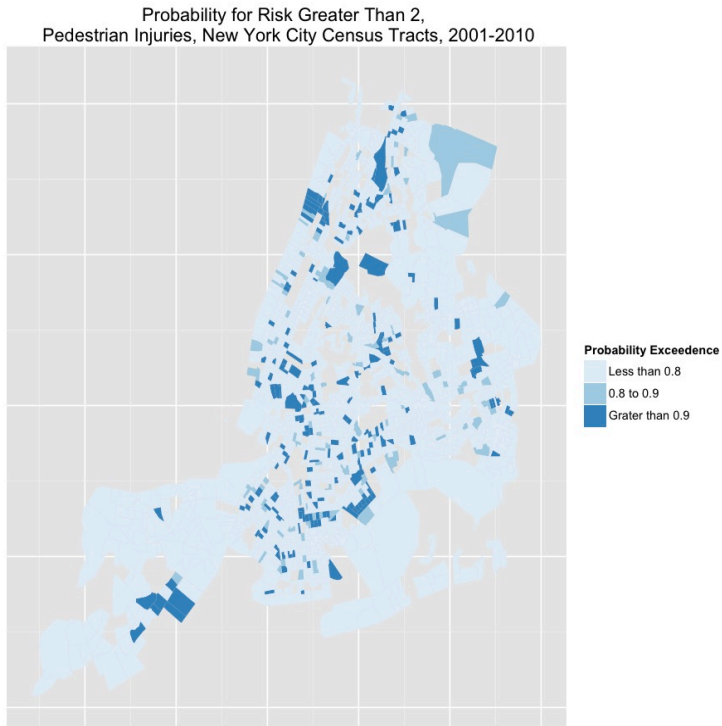


Figure 3: Probability of Incidence Density Ratios Greater Than Two, Spatiotemporal Model of Pedestrian Injury Risk, New York City Census Tracts, 2001-2010

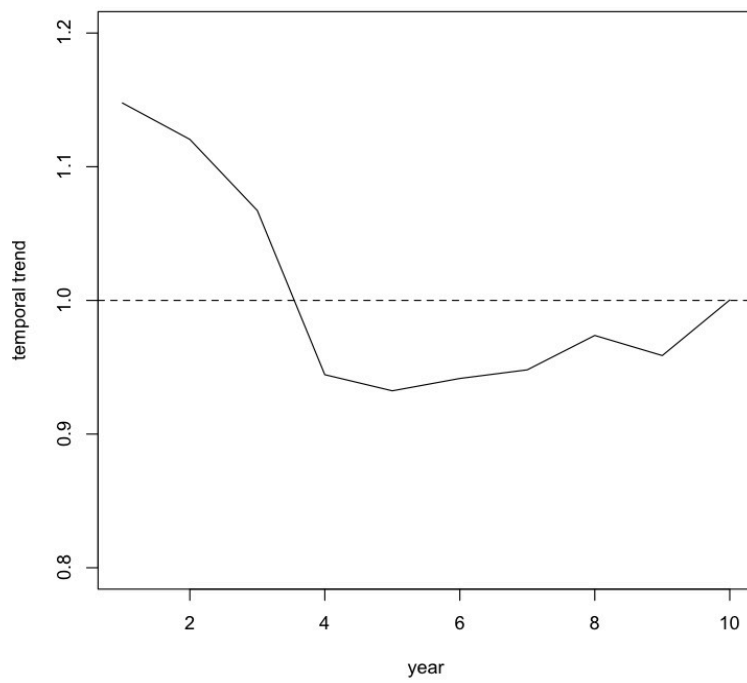
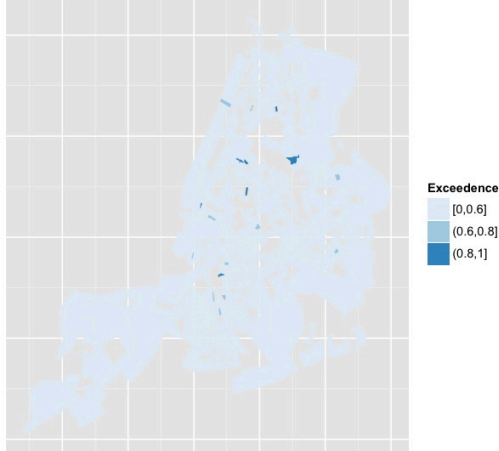
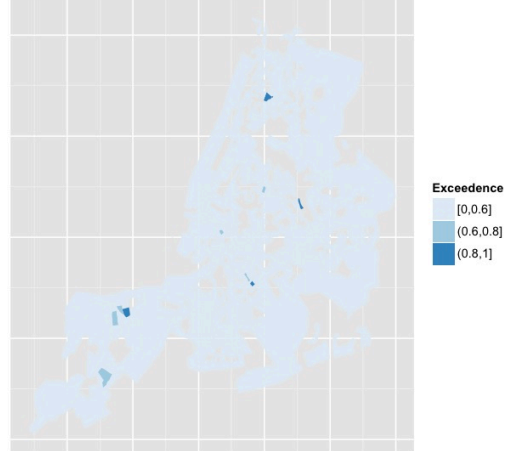


Figure 4: Temporal Trend Term, Spatiotemporal Model of Pedestrian Injury Risk, New York City, 2001-2010

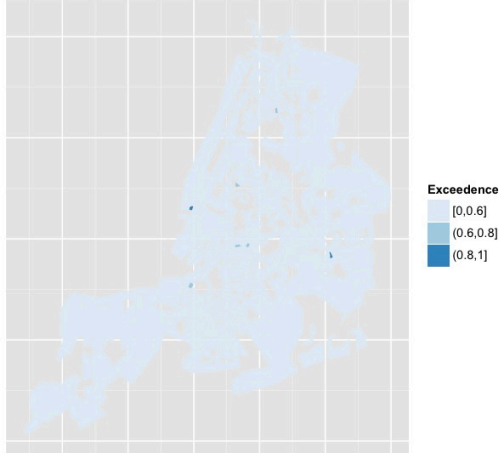
Probability Exceedence RR > 2,  
Pedestrian Injuries, New York City Census Tracts, 2001



Probability Exceedence RR > 2,  
Pedestrian Injuries, New York City Census Tracts, 2007



Probability Exceedence RR > 2,  
Pedestrian Injuries, New York City Census Tracts, 2004



Probability Exceedence RR > 2,  
Pedestrian Injuries, New York City Census Tracts, 2010

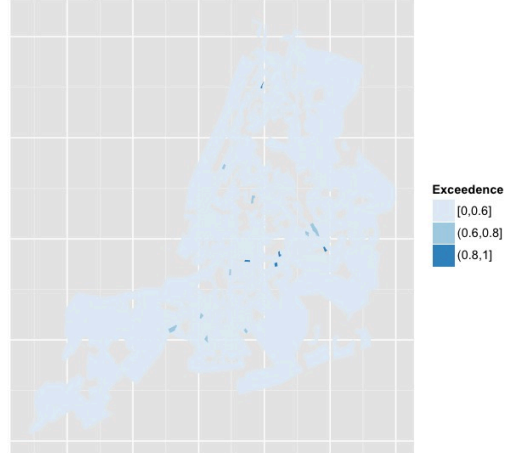


Figure 5: Probability of Incidence Density Ratios Greater Than Two, Space-Time Interaction Term, Spatiotemporal Model of Pedestrian Injury Risk, New York City Census Tracts, 2001-2010

## References

1. Patek G, Thoma T. Commentary: Pedestrian fatalities-a problem on the rise. *Ann Emerg Med.* 2013;62(6):613-615.
2. National Highway Traffic and Safety Administration. US Department of Transportation. Traffic safety facts, 2011 data: bicyclists and other cyclists. 2013; <http://www-nrd.nhtsa.dot.gov/Pubs/811743.pdf> Accessed 18 July 2014.
3. Federal Highway Administration. US Department of Transportation. Pedestrian Safety Strategic Plan: Background Report. 2010; [http://safety.fhwa.dot.gov/ped\\_bike/pssp/background/psafety.cfm](http://safety.fhwa.dot.gov/ped_bike/pssp/background/psafety.cfm) Accessed 18 July 2014.
4. National Highway Traffic and Safety Administration. US Department of Transportation. Traffic Safety Facts, 2007 Data. 2008; [www-nrd.nhtsa.dot.gov/Pubs/810987.pdf](http://www-nrd.nhtsa.dot.gov/Pubs/810987.pdf); Accessed 18 July 2014.
5. Chang D. National Pedestrian Crash Report. National Highway Traffic and Safety Administration. US Department of Transportation. <http://www.walkinginfo.org/library/details.cfm?id=4396>; Accessed 18 July 2014.
6. Viola R, Roe M, Shin H. The New York City Pedestrian Safety Study and Action Plan. 2010; New York City Department of Transportation. [http://www.nyc.gov/html/dot/downloads/pdf/nyc\\_ped\\_safety\\_study\\_action\\_plan.pdf](http://www.nyc.gov/html/dot/downloads/pdf/nyc_ped_safety_study_action_plan.pdf); Accessed 18 July 2014.
7. Office of the Mayor. City of New York. Vision Zero Action Plan 2014; <http://www.nyc.gov/html/visionzero/pdf/nyc-vision-zero-action-plan.pdf>; Accessed 18 July 2014.
8. Vision Zero Initiative. Government of Sweden <http://www.visionzeroinitiative.com/>; Accessed 18 July 2014.
9. Besag J, York J, Mollie A. Bayesian image restoration, with two applications in spatial statistics. *Ann Inst Statist Math.* 1991;43(1):1-59.



10. Blangiardo M, Cameletti M, Baio G, Rue H. Spatial and spatio-temporal models with R-INLA. *Spat Spatiotemporal Epidemiol.* 2013;7:39--55.
11. Rue H, Martino S, Lindgren F, Simpson D, Riebler A. INLA: Functions which allow to perform full Bayesian analysis of latent Gaussian models using Integrated Nested Laplace Approximation. 2013. <http://www.r-inla.org/>; Accessed 18 July 2014.
12. The R Development Team. R: A Language and Environment for Statistical Computing Version 3.0.2. 2014; <http://www.r-project.org/>; Accessed 18 July 2014.
13. Lewin-Koh R, Nicholas J, Bivand. maptools: Tools for reading and handling spatial objects. R package version 08-14. 2012.
14. United States Census Bureau. American factFinder. Department of Commerce. Washington, DC. <http://factfinder2.census.gov/>; Accessed 18 July 2014.
15. Congdon P. Assessing the impact of socioeconomic variables on small area variations in suicide outcomes in England. *Int J Environ Res Public Health.* 2013;10(1):158--177.
16. Pabayo R, Molnar B, Street N, Kawachi I. The relationship between social fragmentation and sleep among adolescents living in Boston, Massachusetts. *J Public Health (Oxf).* 2014 Feb 3. [Epub ahead of print].
17. Ross Z, Ito K, Johnson S, Yee M, Pezeshki G, Clougherty J. Spatial and temporal estimation of air pollutants in New York City: exposure assignment for use in a birth outcomes study. *Environ Health.* 2013;12:51.
18. Karmin C. Hotels Swell in Manhattan. *The Wall Street Journal.* 2013; Commercial Real Estate. <http://online.wsj.com/news/articles/SB10001424052702304441404579121911381744056>; Accessed 18 July 2014.
19. Bivand R. spdep: Spatial dependence: weighting schemes, statistics and models. 2013.
20. Lawson A, Biggeri A, Boehning D, Lesaffre E, Viel J, Clark A, et al. Disease mapping models:

an empirical evaluation. Disease Mapping Collaborative Group. *Stat Med.* 2000;19(17-18):2217-2241.

21. Lawson AB. Bayesian Disease Mapping: Hierarchical Modeling in Spatial Epidemiology, Second Edition (Chapman & Hall/CRC Interdisciplinary Statistics). Boca Raton, Fla. Chapman and Hall/CRC; 2013.

22. Banerjee S, Carlin BP, Gelfand AE, ed. Hierarchical Modeling and Analysis for Spatial Data. Boca Raton, Fla.: Chapman & Hall/CRC; 2004.

23. Bernardinelli L, Clayton D, Pascutto C, Montomoli C, Ghislandi M, Songini M. Bayesian analysis of space-time variation in disease risk. *Stat Med.* 1995;14(21-22):2433-2443.

24. Knorr-Held L. Bayesian modeling of inseparable space-time variation in disease risk. *Stat Med.* 2000;19(17-18):2555-2567.

25. Schrodle B, Held L. A primer on disease mapping and ecological regression using INLA. *Computational Statistics.* 2011;26(2):241-258.

26. Clayton D, Bernardinelli L. Bayesian methods for mapping disease risk. In Elliott P, Cuzick J, English D, Stern R eds. Geographical and Environmental Epidemiology: Methods for Small-Area Studies. Oxford: Oxford University Press, 1992: 205–20.

27. Richardson S, Thomson A, Best N, Elliott P. Interpreting posterior relative risk estimates in disease-mapping studies. *Environ Health Perspect.* 2004;112(9):1016--1025.

28. Best N, Richardson S, Thomson A. A comparison of Bayesian spatial models for disease mapping. *Stat Methods Med Res.* 2005;14(1):35--59

29. DiMaggio C, Li G. Effectiveness of a safe routes to school program in preventing school-aged pedestrian injury. *Pediatrics.* 2013; 131(2):290-296.

30. Link B, Phelan J. Social conditions as fundamental causes of disease. *J Health and Social Behavior.* 1995; extra issue:80-94.

31. Galea S, Tracy M, Hoggatt K, DiMaggio C, Karpati A. Estimated deaths attributable to social factors in the United States. *Am J Public Health*. 2011; 101(8):1456-1465.
32. Allardyce J, Gilmour H, Atkinson J, Rapson T, Bishop J, McCreadie R. Social fragmentation, deprivation and urbanicity: relation to first-admission rates for psychoses. *British J oPsychiatry*. 2005;187(5):401--406.
33. Stjarne MK, de Leon AP, Hallqvist J. Contextual effects of social fragmentation and material deprivation on risk of myocardial infarction: results from the Stockholm Heart Epidemiology Program (SHEEP). *Int J of Epidemiol*. 2004;33(4):732--741.
34. Pabayo R, Barnett TA, Datta GD, Lambert M, O'Loughlin J, Kawachi I. Area-level social fragmentation and walking for exercise: cross-sectional findings from the Quebec Adipose and Lifestyle Investigation in Youth Study. *Am J Public Health*. 2012;102(9):e30-- e37.
35. Sampson RJ, Raudenbush SW, Earls F. Neighborhoods and violent crime: a multilevel study of collective efficacy. *Science*. 1997;277(5328):918--924.
36. DiMaggio C, Galea S, Vlahov D. Bayesian hierarchical spatial modeling of substance abuse patterns following a mass trauma: the role of time and place. *Subst Use Misuse*. 2009;44(12):1725-1743.
37. Susser E. Eco-epidemiology: thinking outside the black box. *Epidemiology*. 2004;15(5):519-20; author reply 527-8.
38. Wang Y, Nomura Y, Pat-Horenczyk R, Doppelt O, Abramovitz R, Brom D, et al. Association of direct exposure to terrorism, media exposure to terrorism, and other trauma with emotional and behavioral problems in preschool children. *Ann N Y Acad Sci*. 2006;1094:363-368.
39. Ben-Shlomo Y. Rising to the challenges and opportunities of life course epidemiology. *Int J Epidemiol*. 2007;36(3):481-483.
40. Galea S, Nandi A, Vlahov D. The social epidemiology of substance use. *Epidemiol Rev*.

2004;26:36-52.

41. Thomas AJ, Eberly LE, Davey Smith G, Neaton JD. ZIP-code-based versus tract-based income measures as long-term risk-adjusted mortality predictors. *Am J Epidemiol.* 2006;164(6):586-590.
42. Richardson D, Volkow N, Kwan M, Kaplan R, Goodchild M, Croyle R. Medicine. Spatial turn in health research. *Science.* 2013; 339(6126):1390-1392.
43. Hu S, Ivan JN, Ravishanker N, Mooradian J. Temporal modeling of highway crash counts for senior and non-senior drivers. *Accident Analysis & Prevention.* 2013;50:1003-1013.
44. Miaou SP, Song JJ, Mallick BK. Roadway traffic crash mapping: a space-time modeling approach. *Journal of Transportation and Statistics.* 2003;6:33-58.
45. Wang Y, Kockelman KM. A Poisson-lognormal conditional-autoregressive model for multivariate spatial analysis of pedestrian crash counts across neighborhoods. *Accident Analysis & Prevention.* 2013;60:71-84.
46. Baath R. Easy Laplace Approximation of Bayesian Models in R. <http://www.sumsar.net/blog/2013/11/easy-laplace-approximation/>; Accessed 18 July 2014.
47. Hoffman MD, Gelman A. The no-U-turn sampler: Adaptively setting path lengths in Hamiltonian Monte Carlo. J arXiv preprint arXiv:11114246. 2011;. <http://www.stat.columbia.edu/~gelman/research/published/nuts.pdf>; Accessed 18 July 2014.
48. King M. UK homeownership falls to lowest level since 1988; The Guardian, 15 November 2012 <http://www.theguardian.com/money/2012/nov/16/home-ownership-lowest-since-1988> Accessed 15 August 2015.
49. Retting RA. Urban pedestrian safety. *Bull N Y Acad Med.* 1988; 64(7):810-815.
50. Greenland S, Schlesselman J, Criqui M. The fallacy of employing standardized regression coefficients and correlations as measures of effect. *Am J Epidemiol.* 1986;123(2):203--208.
51. Newman T, Browner W. In defense of standardized regression coefficients. *Epidemiology.*

1991;2(5):383--386.

52. Cottrill C, Thakuria P. Evaluating pedestrian crashes in areas with high low-income or minority populations. *Accident Analysis & Prevention*. 2010; 42(6):1718-1728.

53. Hameed S, Popkin C, Cohn S, Johnson E. The epidemic of pediatric traffic injuries in South Florida: a review of the problem and initial results of a prospective surveillance strategy. *Am J Public Health*. 2004; 94(4):554--556.

55. Rundle A, Bader M, Richards C, Neckerman K, Teitler J. Using Google Street View to audit neighborhood environments. *Am J Prev Med*. 2011;40(1):94-100.

55. Mooney S, Rundle A. Open CANVAS: A Web Application Harnessing Google Street View to Collect and Share Data on Neighborhoods. 2013;. <https://www.newschallenge.org/challenge/healthdata/refinement/canvas-using-google-street-view-to-gather-and-share-data-on-neigh>  
Accessed 18 July 2014.

56. Hossain M, Lawson AB. Cluster detection diagnostics for small area health data: with reference to evaluation of local likelihood models. *Statistics in Medicine*. 2006;25(5):771-786.

57. Hossain MM, Lawson AB. Space-time Bayesian small area disease risk models: development and evaluation with a focus on cluster detection. *Environmental and ecological statistics*. 2010;17(1):73--95.

58. Lawson A. Comment: Exceedence probabilities in the analysis of disease maps where risk anomalies are present. 2006;. [http://www.sph.sc.edu/alawson/Web\\_comment\\_exceedence.pdf](http://www.sph.sc.edu/alawson/Web_comment_exceedence.pdf); Accessed 18 July 2014.